manner. Consider a family of planes tangent to S along the curve c. This family determines a surface E, enveloping these tangent planes, which possesses the property that it will be tangent to S along the curve c and whose Gaussian curvature $K \equiv 0$. (Cf. M. do Carmo [dC 2] pp. 195–197). It is not difficult to show that the parallelism along c, defined through the vanishing of the covariant derivative is the same whether considered relative to S or relative to E. On the other hand, surfaces of zero curvature can be shown to be locally isometric to a plane. Since parallelism is invariant by isometry, we can perform it "euclideanly" in the isometric image of E and then bring it back to S. This was the construction used classically to define parallelism. (M. do Carmo [dC 2] p. 244). It will turn out that it is preferable, technically, to work with the covariant derivative.

The notion of covariant derivative has many important consequences. It makes it clear that the two basic ideas of geodesic and curvature can be defined in more general situations than that of Riemannian manifolds. To this end it suffices that one be able to define a notion of derivation of vector fields with certain properties (which nowadays we call an affine connection, Cf. Definition 2.1 of this chapter). This has stimulated the creation of many different "geometric structures" (on differentiable manifolds) more general than Riemannian geometry. In the same way as metric Euclidean geometry is a particular case of affine geometry and more generally of projective geometry, Riemannian geometry is a particular case of more general geometric structures.

We are not going to enter into the details of these developments. Our interest in affine connections rests in the fact (Cf. Theorem 3.6 of this chapter) that a choice of a Riemannian metric on a manifold M uniquely determines a certain affine connection on M. We are then able, in this fashion, to differentiate vector fields on M.

2. Affine Connections

Let us indicate by $\mathcal{X}(M)$ the set of all vector fields of class C^∞ on M and by $\mathcal{D}(M)$ the ring of real-valued functions of class C^∞ defined on M.

2.1 DEFINITION. An *affine connection* ∇ on a differentiable manifold M is a mapping

$$\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$$

which is denoted by $(X,Y) \xrightarrow{\nabla} \nabla_X Y$ and which satisfies the following properties :

i) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z.$
ii) $\nabla_X(Y+Z) = \nabla_X Y + \nabla_X Z.$
iii) $\nabla_X(fY) = f\nabla_X Y + X(f)Y,$

in which $X, Y, Z \in \mathcal{X}(M)$ and $f, g \in \mathcal{D}(M)$.

This definition is not as transparent as that of Riemannian structure. The following proposition, nevertheless, should clarify the situation a little.

2.2 PROPOSITION. *Let M be a differentiable manifold with an affine connection ∇. There exists a unique correspondence which associates to a vector field V along the differentiable curve $c: I \to M$ another vector field $\frac{DV}{dt}$ along c, called the covariant derivative of V along c, such that:*

a) $\frac{D}{dt}(V+W) = \frac{DV}{dt} + \frac{DW}{dt}.$

b) $\frac{D}{dt}(fV) = \frac{df}{dt}V + f\frac{DV}{dt}$*, where W is a vector field along c and f is a differentiable function on I.*

c) If V is induced by a vector field $Y \in \mathcal{X}(M)$, i.e., $V(t) = Y(c(t))$, then $\frac{DV}{dt} = \nabla_{dc/dt} Y$.

2.3 REMARK. The last line of (c) makes sense, since $\nabla_X Y(p)$ depends on the value of $X(p)$ and the value Y along a curve, tangent to X at p. In effect, part (iii) of Definition 2.1 allows us to show that the notion of affine connection is actually a local notion (cf. Rem. 5.7 of Chap. 0). Choosing a system of coordinates $(x_1, \ldots, x_n)$ about p and writing

$$X = \sum_i x_i X_i, \qquad Y = \sum_j y_j Xj,$$

where $X_i = \frac{\partial}{\partial x_i}$, we have

$$\nabla_X Y = \sum_i x_i \nabla_{X_i}\Big(\sum_j y_j X_j\Big) = \sum_{ij} x_i y_j \nabla_{X_i} X_j + \sum_{ij} x_i X_i(y_j) X_j.$$

Setting $\nabla_{X_i} X_j = \sum_k \Gamma_{ij}^k X_k$, we conclude that the Γ_{ij}^k are differentiable functions and that

$$\nabla_X Y = \sum_k \Big(\sum_{ij} x_i y_j \Gamma_{ij}^k + X(y_k)\Big) X_k,$$

which proves that $\nabla_X Y(p)$ depends on $x_i(p)$, $y_k(p)$ and the derivatives $X(y_k)(p)$ of y_k by X.

2.4 Remark. The proposition above shows that the choice of an affine connection on M leads to a bona fide (i.e. satisfying (a) and (b)) derivative of vector fields along curves. The notion of connection furnishes, therefore, a manner of differentiating vectors along curves; in particular, it will then be possible to speak of the acceleration of a curve in M.

Proof of Proposition 2.2. Let us suppose initially that there exists a correspondence satisfying (a), (b) and (c). Let $\mathbf{x}: U \subset \mathbf{R}^n \to M$ be a system of coordinates with $c(I) \cap \mathbf{x}(U) \neq \phi$ and let $(x_1(t), x_2(t), \dots, x_n(t))$ be the local expression of $c(t)$, $t \in I$. Let $X_i = \frac{\partial}{\partial x_i}$. Then we can express the field V locally as $V = \sum_j v^j X_j$, $j = 1, \dots, n$, where $v^j = v^j(t)$ and $X_j = X_j(c(t))$.

By a) and b), we have

$$\frac{DV}{dt} = \sum_j \frac{dv^j}{dt} X_j + \sum_j v^j \frac{DX_j}{dt}.$$

By c) and (i) of Definition 2.1,

$$\begin{aligned}\frac{DX_j}{dt} &= \nabla_{dc/dt} X_j = \nabla_{(\Sigma \frac{dx_i}{dt} X_i)} X_j \\ &= \sum_i \frac{dx_i}{dt} \nabla_{X_i} X_j, \quad i, j = 1, \dots, n.\end{aligned}$$

Therefore,

$$\text{(1)} \qquad \frac{DV}{dt} = \sum_j \frac{dv^j}{dt} X_j + \sum_{i,j} \frac{dx_i}{dt} v^j \nabla_{X_i} X_j.$$

The expression (1) shows us that if there is a correspondence satisfying the conditions of Proposition 2.2, then such a correspondence is unique.

To show existence, define $\frac{DV}{dt}$ in $\mathbf{x}(U)$ by (1). It is easy to verify that (1) possesses the desired properties. If $\mathbf{y}(W)$ is another coordinate neighborhood, with $\mathbf{y}(W)\cap\mathbf{x}(U) \neq \phi$ and we define $\frac{DV}{dt}$ in $\mathbf{y}(W)$ by (1), the definitions agree in $\mathbf{y}(W)\cap\mathbf{x}(U)$, by the uniqueness of $\frac{DV}{dt}$ in $\mathbf{x}(U)$. It follows that the definition can be extended over all of M, and this concludes the proof. $\square$

The concept of parallelism now follows in a natural manner.

2.5 DEFINITION. Let M be a differentiable manifold with an affine connection ∇. A vector field V along a curve $c\colon I \to M$ is called *parallel* when $\frac{DV}{dt} = 0$, for all $t \in I$.

2.6 PROPOSITION. *Let M be a differentiable manifold with an affine connection ∇. Let $c\colon I \to M$ be a differentiable curve in M and let V_o be a vector tangent to M at $c(t_o)$, $t_o \in I$ (i.e. $V_o \in T_{c(t_o)}M$). Then there exists a unique parallel vector field V along c, such that $V(t_o) = V_o$, (($V(t)$ is called the parallel transport of $V(t_o)$ along c).*

Proof. Suppose that the theorem was proved for the case in which $c(I)$ is contained in a local coordinate neighborhood. By compactness, for any $t_1 \in I$, the segment $c([t_o, t_1]) \subset M$ can be covered by a finite number of coordinate neighborhoods, in each of which V can be defined, by hypothesis. From uniqueness, the definitions coincide when the intersections are not empty, thus allowing the definition of V along all of $[t_o, t_1]$.

We have only, therefore, to prove the theorem when $c(I)$ is contained in a coordinate neighborhood $\mathbf{x}(U)$ of a system of coordinates $\mathbf{x}\colon U \subset \mathbf{R}^n \to M$. Let $\mathbf{x}^{-1}(c(t)) = (x_1(t), \dots, x_n(t))$ be the local expression for $c(t)$ and let $V_o = \sum_j v_o^j X_j$, where $X_j = \frac{\partial}{\partial x_j}(c(t_o))$.

Suppose that there exists a vector field V in $\mathbf{x}(U)$ which is parallel along c with $V(t_o) = V_o$. Then $V = \sum v^j X_j$ satisfies

$$0 = \frac{DV}{dt} = \sum_j \frac{dv^j}{dt} X_j + \sum_{i,j} \frac{dx_i}{dt} v^j \nabla_{X_i} X_j.$$

Putting $\nabla_{X_i} X_j = \sum_k \Gamma_{ij}^k X_k$, and replacing j with k in the first sum, we obtain

$$\frac{DV}{dt} = \sum_k \left\{ \frac{dv^k}{dt} + \sum_{i,j} v^j \frac{dx_i}{dt} \Gamma_{ij}^k \right\} X_k = 0.$$

The system of n differential equations in $v^k(t)$,

$$(2) \qquad 0 = \frac{dv^k}{dt} + \sum_{i,j} \Gamma^k_{ij} v^j \frac{dx_i}{dt}, \quad k = 1, \ldots, n,$$

possesses a unique solution satisfying the initial conditions $v^k(t_o) = v^k_o$. It then follows that, if V exists, it is unique. Moreover, since the system is linear, any solution is defined for all $t \in I$, which then proves the existence (and uniqueness) of V with the desired properties. $\square$

3. Riemannian Connections

3.1 Definition. Let M be a differentiable manifold with an affine connection ∇ and a Riemannian metric $\langle\, ,\, \rangle$. A connection is said to be *compatible* with the metric $\langle\, ,\, \rangle$, when for any smooth curve c and any pair of parallel vector fields P and P' along c, we have $\langle P, P' \rangle = \text{constant}$.

Definition 3.1 is justified by the following proposition which shows that if ∇ is compatible with $\langle\, ,\, \rangle$, then we are able to differentiate the inner product by the usual "product rule".

3.2 Proposition. *Let M be a Riemannian manifold. A connection ∇ on M is compatible with a metric if and only if for any vector fields V and W along the differentiable curve $c\colon I \to M$ we have*

$$(3) \qquad \frac{d}{dt}\langle V, W \rangle = \langle \frac{DV}{dt}, W \rangle + \langle V, \frac{DW}{dt} \rangle, \quad t \in I.$$

Proof. It is obvious that equation (3) implies that ∇ is compatible with $\langle\, ,\, \rangle$. Therefore, let us prove the converse. Choose an orthonormal basis $\{P_1(t_o), \ldots, P_n(t_o)\}$ of $T_{x(t_o)}(M)$, $t_o \in I$. Using Proposition 2.6, we can extend the vectors $P_i(t_o)$, $i = 1, \ldots, n$, along c by parallel translation. Because ∇ is compatible with the metric, $\{P_1(t), \ldots, P_n(t)\}$ is an orthonormal basis of $T_{c(t)}(M)$, for any $t \in I$. Therefore, we can write

$$V = \sum_i v^i P_i, \qquad W = \sum_i w^i P_i, \qquad i = 1, \ldots, n$$

where v^i and w^i are differentiable functions on I. It follows that

$$\frac{DV}{dt} = \sum_i \frac{dv^i}{dt} P_i, \qquad \frac{DW}{dt} = \sum_i \frac{dw^i}{dt} P_i.$$

Therefore,

$$\begin{aligned}\langle \frac{DV}{dt}, W\rangle + \langle V, \frac{DW}{dt}\rangle &= \sum_i \left\{ \frac{dv^i}{dt} w^i + \frac{dw^i}{dt} v^i \right\} \\ &= \frac{d}{dt}\left\{ \sum_i v^i w^i \right\} = \frac{d}{dt}\langle V, W\rangle. \quad \square\end{aligned}$$

3.3 COROLLARY. *A connection ∇ on a Riemannian manifold M is compatible with the metric if and only if*

(4) $\quad X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle, \qquad X, Y, Z \in \mathcal{X}(M).$

Proof. Suppose that ∇ is compatible with the metric. Let $p \in M$ and let $c: I \to M$ be a differentiable curve with $c(t_o) = p$, $t_o \in I$, and with $\frac{dc}{dt}\big|_{t=t_o} = X(p)$. Then

$$X(p)\langle Y, Z\rangle = \frac{d}{dt}\langle Y, Z\rangle\Big|_{t=t_o} = \langle \nabla_{X(p)} Y, Z\rangle_p + \langle Y, \nabla_{X(p)} Z\rangle_p.$$

Since p is arbitrary, (4) follows. The converse is obvious. $\square$

3.4 DEFINITION. An affine connection ∇ on a smooth manifold M is said to be *symmetric* when

(5) $\quad \nabla_X Y - \nabla_Y X = [X, Y] \quad$ for all $\quad X, Y \in \mathcal{X}(M).$

3.5 REMARK. In a coordinate system $(U, \mathbf{x})$, the fact that ∇ is symmetric implies that for all $i, j = 1, \ldots, n$,

(5') $\quad \nabla_{X_i} X_j - \nabla_{X_j} X_i = [X_i, X_j] = 0, \quad X_i = \dfrac{\partial}{\partial x_i},$

which justifies the terminology (observe that (5') is equivalent to the fact that $\Gamma^k_{ij} = \Gamma^k_{ji}$).

We are now able to state the fundamental theorem of this chapter.

3.6 Theorem. *(Levi-Civita). Given a Riemannian manifold M, there exists a unique affine connection ∇ on M satisfying the conditions:*

a) ∇ is symmetric.
b) ∇ is compatible with the Riemannian metric.

Proof. Suppose initially the existence of such a ∇. Then

$$X\langle Y,Z\rangle = \langle\nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle, \tag{6}$$

$$Y\langle Z,X\rangle = \langle\nabla_Y Z, X\rangle + \langle Z, \nabla_Y X\rangle, \tag{7}$$

$$Z\langle X,Y\rangle = \langle\nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle. \tag{8}$$

Adding (6) and (7) and subtracting (8), we have, using the symmetry of ∇, that

$$\begin{aligned} &X\langle Y,Z\rangle + Y\langle Z,X\rangle - Z\langle X,Y\rangle \\ &\qquad = \langle [X,Z],Y\rangle + \langle [Y,Z],X\rangle + \langle [X,Y],Z\rangle + 2\langle Z,\nabla_Y X\rangle. \end{aligned}$$

Therefore

$$\begin{aligned} \langle Z,\nabla_Y X\rangle = \frac{1}{2}\{&X\langle Y,Z\rangle + Y\langle Z,X\rangle - Z\langle X,Y\rangle \\ &-\langle [X,Z],Y\rangle - \langle [Y,Z],X\rangle - \langle [X,Y],Z\rangle\}. \end{aligned} \tag{9}$$

The expression (9) shows that ∇ is uniquely determined from the metric $\langle\ ,\ \rangle$. Hence, if it exists, it will be unique.

To prove existence, define ∇ by (9). It is easy to verify that ∇ is well-defined and that it satisfies the desired conditions. $\square$

3.7 Remark. The connection given by the theorem will be referred to, from now on, as the *Levi-Civita* (or *Riemannian*) *connection* on M.

Let us conclude this chapter by writing part of what was shown above in a coordinate system $(U,\mathbf{x})$. It is customary to call the functions Γ_{ij}^k defined on U by $\nabla_{X_i}X_j = \sum_k \Gamma_{ij}^k X_k$, the *coefficients of the connection* ∇ on U or the *Christoffel symbols* of the connection. From (9) it follows that

$$\sum_\ell \Gamma_{ij}^\ell g_{\ell k} = \frac{1}{2}\left\{\frac{\partial}{\partial x_i}g_{jk} + \frac{\partial}{\partial x_j}g_{ki} - \frac{\partial}{\partial x_k}g_{ij}\right\},$$

where $g_{ij} = \langle X_i, X_j \rangle$.

Since the matrix (g_{km}) admits an inverse (g^{km}), we obtain that

$$(10) \qquad \Gamma_{ij}^m = \frac{1}{2} \sum_k \left\{ \frac{\partial}{\partial x_i} g_{jk} + \frac{\partial}{\partial x_j} g_{ki} - \frac{\partial}{\partial x_k} g_{ij} \right\} g^{km}.$$

The equation (10) is a classical expression for the Christoffel symbols of the Riemannian connection in terms of the g_{ij} (given by the metric).

Observe that for the Euclidean space $\mathbf{R}^n$, we have $\Gamma_{ij}^k = 0$.

In terms of the Christoffel symbols, the covariant derivative has the classical expression

$$\frac{DV}{dt} = \sum_k \left\{ \frac{dv^k}{dt} + \sum_{i,j} \Gamma_{ij}^k v^j \frac{dx_i}{dt} \right\} X_k$$

which follows from (1). Observe that $\frac{DV}{dt}$ differs from the usual derivative in Euclidean space by terms which involve the Christoffel symbols. Therefore, in Euclidean spaces the covariant derivative coincides with the usual derivative.

EXERCISES

1. Let M be a Riemannian manifold. Consider the mapping

$$P = P_{c,t_o,t} \colon T_{c(t_o)}M \to T_{c(t)}M$$

defined by: $P_{c,t_o,t}(v)$, $v \in T_{c(t_o)}M$, is the vector obtained by parallel transporting the vector v along the curve c. Show that P is an isometry and that, if M is oriented, P preserves the orientation.

2. Let X and Y be differentiable vector fields on a Riemannian manifold M. Let $p \in M$ and let $c \colon I \to M$ be an integral curve